

## Impact of non-nucleotidic bulges on recognition of mixed-sequence dsDNA by pyrene-functionalized Invader probes.

Dale C. Guenther, Raymond G. Emehiser, Allison Inskeep, Saswata Karmakar and Patrick J.

Hrdlicka.<sup>†,\*</sup>

Department of Chemistry, University of Idaho, Moscow, ID 83844-2343, USA

E-mail: [hrdlicka@uidaho.edu](mailto:hrdlicka@uidaho.edu)

### ELECTRONIC SUPPLEMENTARY INFORMATION

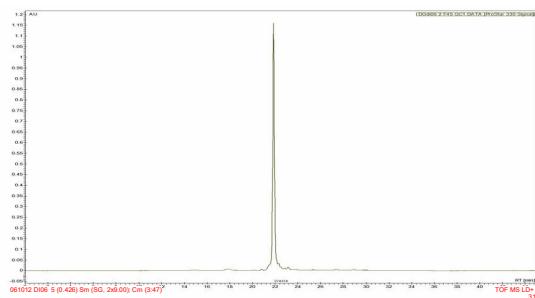
MALDI-MS of <b>ON1-ON14</b> (Table S1)	S2
HPLC traces and MS spectra of <b>ON1-ON14</b> (Figs. S1-S4)	S3-6
Representative thermal denaturation curves (Fig. S5)	S7
Binding specificity of individual probe strands <b>ON1-ON14</b> (Table S2)	S8-9
Thermal advantage values for different Invader probes (Table S3)	S10
Additional discussion of the $\Delta G_{\text{rec}}$ term	S11
Additional discussion of thermodynamic parameters (Tables S4-S8)	S11-20
Representative electrophoretograms for recognition of <b>DH1</b> using different Invader probes at room temperature or 37 °C (Fig. S6)	S21
Degree of <b>DH1</b> -recognition using different Invader probes (Table S9)	S22
Dose-response experiments using select bulge-containing Invader probes (Fig. S7)	S23
Sequences and $T_m$ values of DNA hairpins used in this study (Table S10)	S24
MALDI-MS of <i>DYZ-1</i> targeting ONs (Table S11)	S25
HPLC traces and MS spectra of <i>DYZ-1</i> targeting ONs (Figs. S8-S10)	S26-28
Representative images of <i>DYZ-1</i> targeting Invader probes (Figs. S11-S14)	S29-32
Supplementary references	S33

**Table S1.** MALDI-MS of **ON1-ON14**.<sup>a</sup>

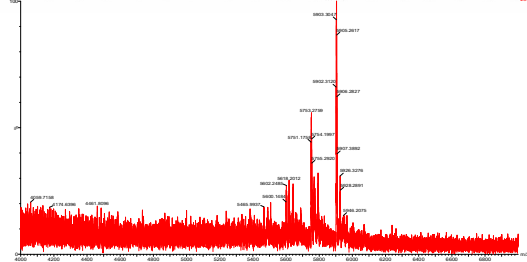
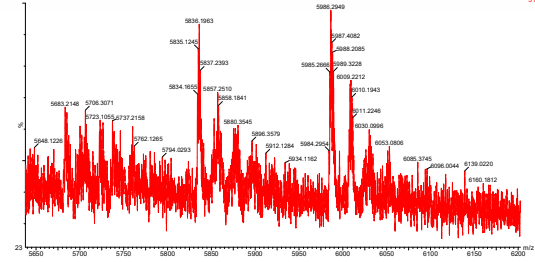
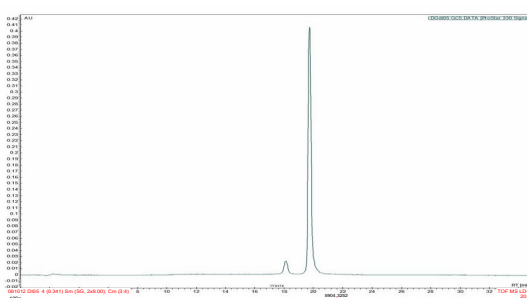
ON	Sequence	Observed	Calculated
		<i>m/z</i> [M+H] <sup>+</sup>	<i>m/z</i> [M+H] <sup>+</sup>
<b>1</b>	5'-GG <u>I</u> GGTCAA <u>C</u> ACTATC <u>I</u> GGA	5986	5984
<b>2</b>	3'-CCA <u>C</u> CAGTTGATAGAC <u>C</u> CT	5904	5901
<b>3</b>	5'-GG <u>I</u> GGTCAA <b>2</b> CTATC <u>I</u> GGA	6139	6137
<b>4</b>	3'-CCA <u>C</u> CAGTT <b>2</b> GATAGAC <u>C</u> CT	6057	6054
<b>5</b>	5'-GG <u>I</u> GGTCAA <b>222</b> CTATC <u>I</u> GGA	6446	6443
<b>6</b>	3'-CCA <u>C</u> CAGTT <b>222</b> GATAGAC <u>C</u> CT	6363	6360
<b>7</b>	5'-GG <u>I</u> GGTCAA <b>4</b> CTATC <u>I</u> GGA	6139	6136
<b>8</b>	3'-CCA <u>C</u> CAGTT <b>4</b> GATAGAC <u>C</u> CT	6055	6053
<b>9</b>	5'-GG <u>I</u> GGTCAA <b>444</b> CTATC <u>I</u> GGA	6443	6440
<b>10</b>	3'-CCA <u>C</u> CAGTT <b>444</b> GATAGAC <u>C</u> CT	6360	6357
<b>11</b>	5'-GG <u>I</u> GGTCAA <b>9</b> CTATC <u>I</u> GGA	6209	6206
<b>12</b>	3'-CCA <u>C</u> CAGTT <b>9</b> GATAGAC <u>C</u> CT	6126	6123
<b>13</b>	5'-GG <u>I</u> GGTCAA <b>999</b> CTATC <u>I</u> GGA	6653	6650
<b>14</b>	3'-CCA <u>C</u> CAGTT <b>999</b> GATAGAC <u>C</u> CT	6570	6567

<sup>a</sup> I/C = 2'-*O*-(pyren-1-yl)methyl-RNA U or C monomers (U = uracil-1-yl; C = cytosin-1-yl); **2** = 1-amino-3-hydroxyprop-2-yl monomer, **4** = 4-hydroxybutyl monomer, **9** = 9-hydroxynonyl monomer. See Figure 1 in the main manuscript for structures. 2,4,6-Trihydroxyacetophenone was used as the matrix.

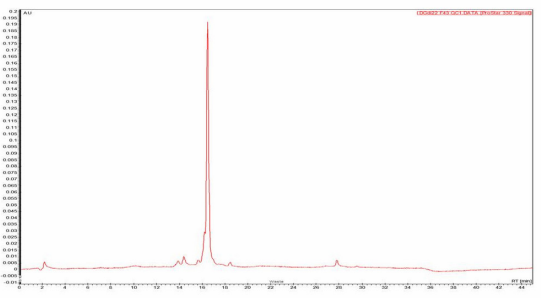
ON1



ON2



ON3



ON4

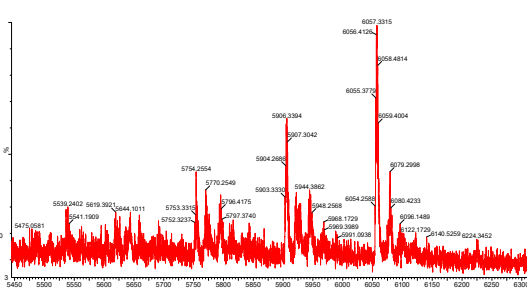
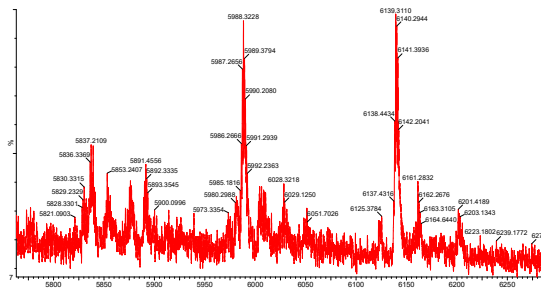
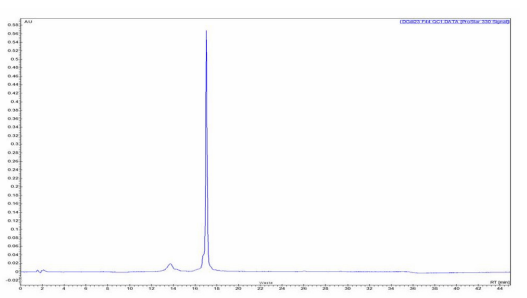
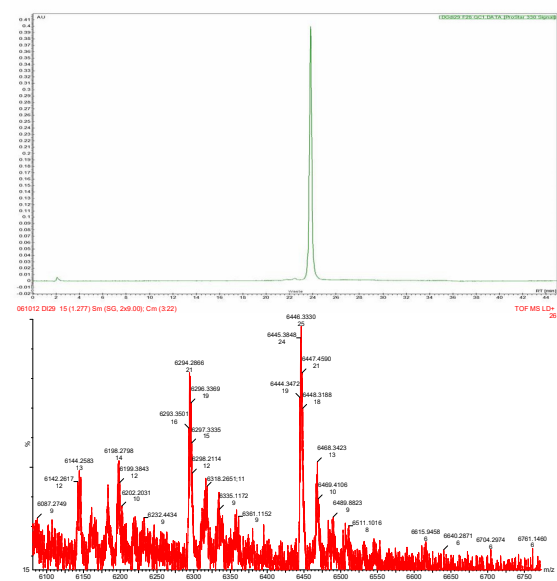
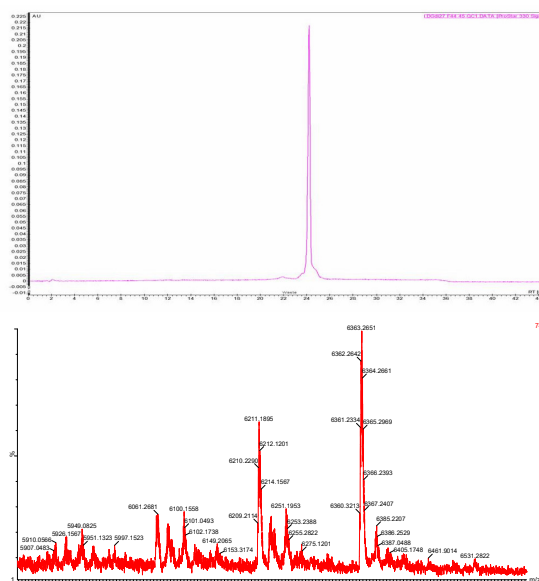


Figure S1. HPLC traces and MS spectra for ON1-ON4.

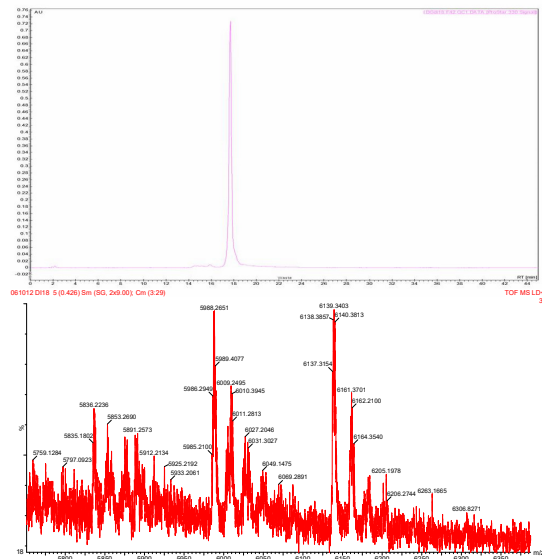
ON5



ON6



ON7



ON8

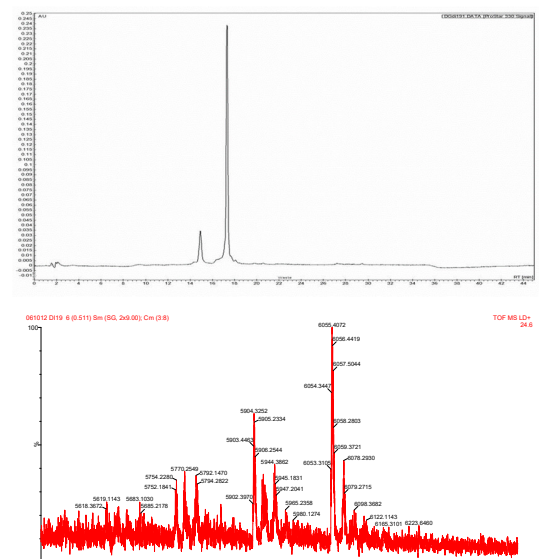
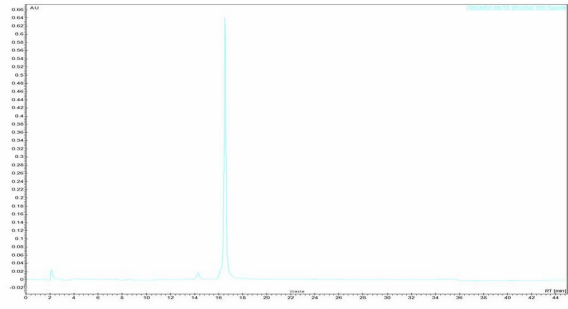
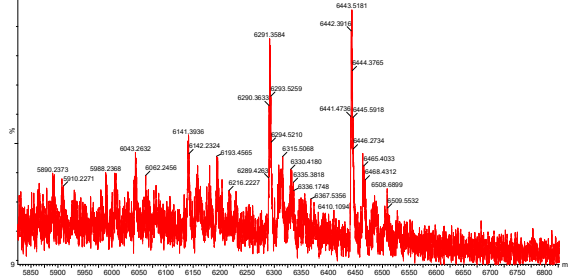


Figure S2. HPLC traces and MS spectra for ON5-ON8.

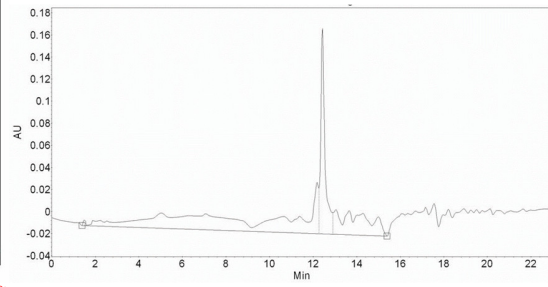
ON9



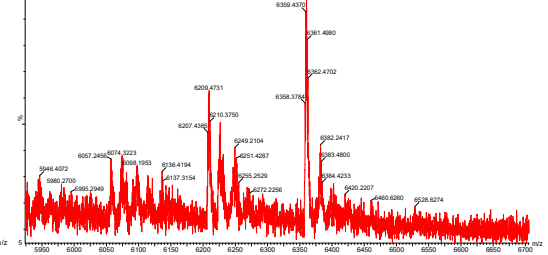
061012 D00 13 (1.107) Sm (SG, 2x9.00) Cm (3.20) TOF MS LD+ 23



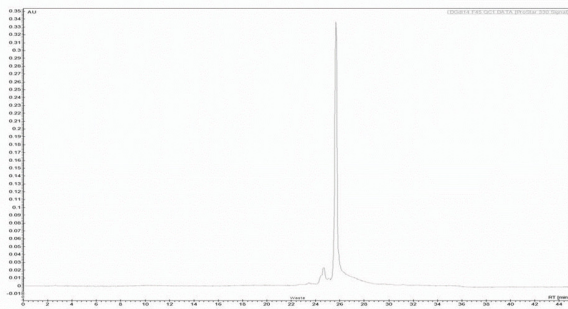
ON10



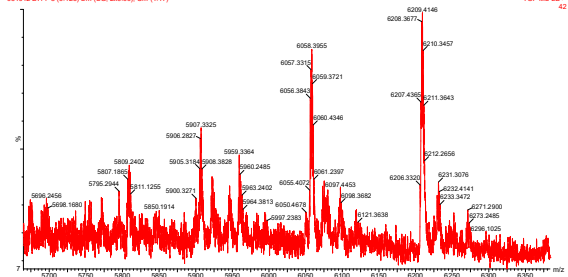
061012 D15 5 (0.420) Sm (SG, 2x9.00) Cm (1.7) TOF MS LD+ 24.6



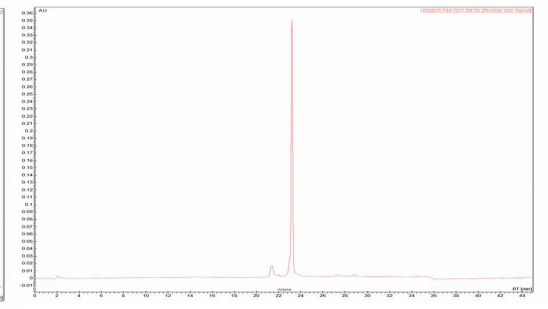
ON11



061012 D14 5 (0.426) Sm (SG, 2x9.00) Cm (1.17) TOF MS LD+ 42



ON12



061012 D15 5 (0.426) Sm (SG, 2x9.00) Cm (1.7) TOF MS LD+ 34.6

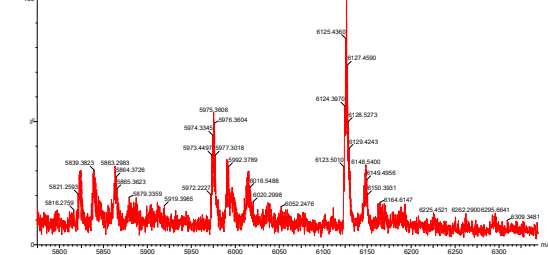


Figure S3. HPLC traces and MS spectra for ON9-ON12.

ON13

ON14

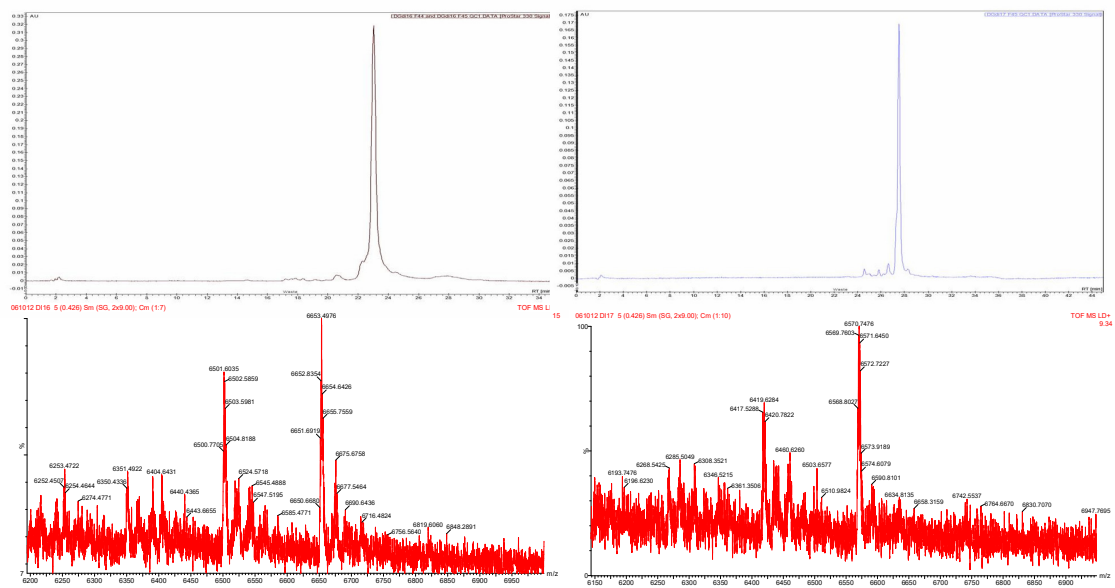
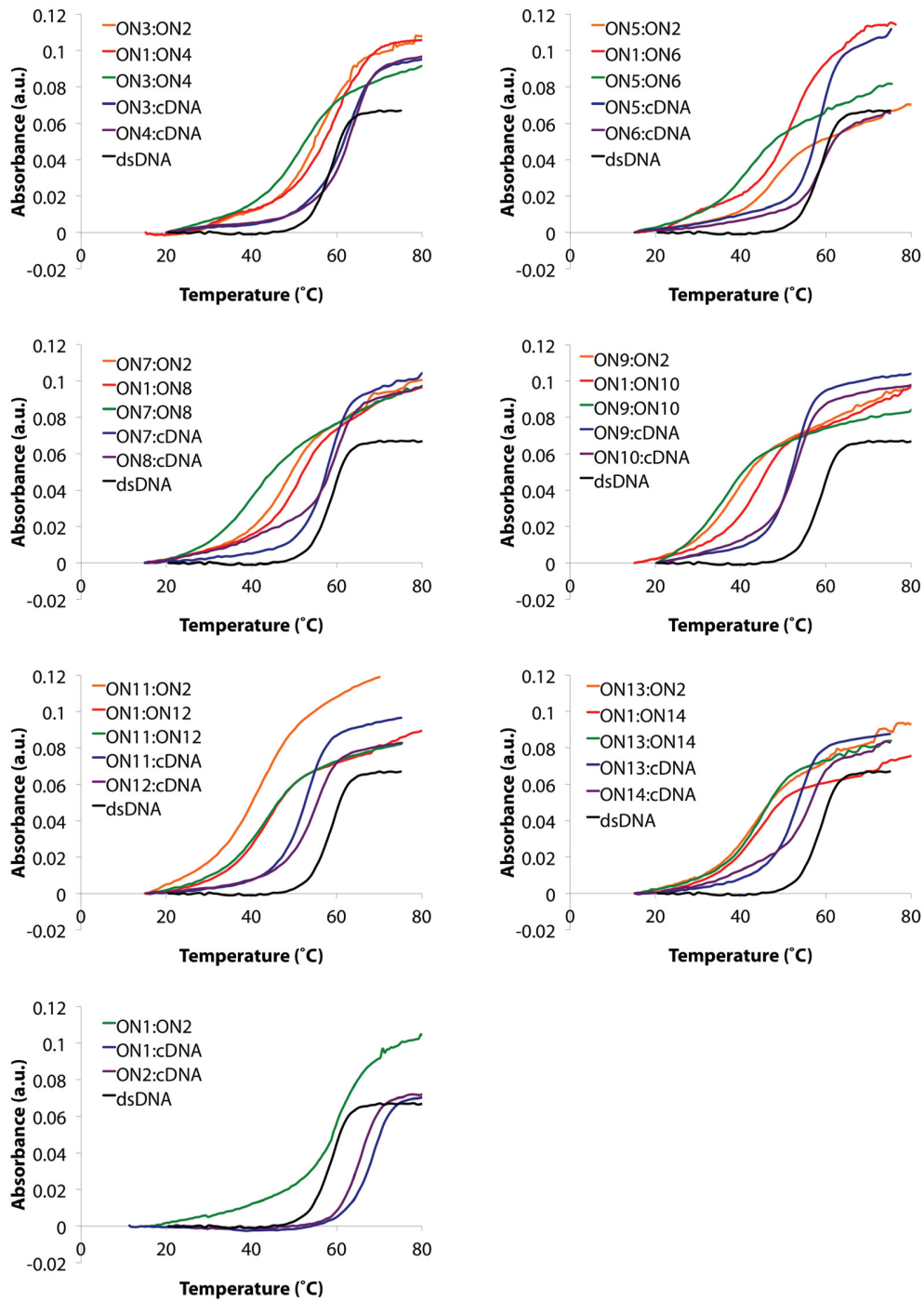


Figure S4. HPLC traces and MS spectra for ON13 and ON14.



**Figure S5.** Representative thermal denaturation curves of double-stranded probes, duplexes between individual probe strands and cDNA, and the unmodified reference DNA duplex (dsDNA). For experimental conditions, see Table 1.

*Binding specificity of individual probe strands ON1-ON14.* The binding specificity of individual probe strands was evaluated using three different singly mismatched single-stranded DNA targets, in which the mismatched nucleotides are located close to the bulge (**MM1** or **MM2**), in the unmodified region between the bulge and the 2'-*O*-(pyren-1-yl)methyl RNA monomer (**MM3** or **MM4**), or close to the 2'-*O*-(pyren-1-yl)methyl RNA monomer (**MM5** or **MM6**) (Table S2).

Conventional probe strands **ON1** and **ON2** discriminate mismatched targets with similar specificity as the corresponding unmodified strands **DNA1** and **DNA2**, except when the mismatched base-pair is located near the 2'-*O*-(pyren-1-yl)methyl RNA monomer, in which case slightly less efficient discrimination is observed (e.g., compare  $\Delta T_{ms}$  for mismatched duplexes entailing **ON1** and **DNA1**). This is in agreement with results from prior studies in which ONs modified with 2'-*O*-(pyren-1-yl)methyluridine monomers were shown to display reduced mismatch discrimination.<sup>S1</sup>

Bulge-containing probe strands display similar or slightly improved discrimination of mismatched targets vis-à-vis conventional probe strands when the mismatched base-pair is located in the unmodified region or opposite of the 2'-*O*-(pyren-1-yl)methyl RNA monomers (e.g., compare  $\Delta T_{ms}$  of **ON12:MM3** with **ON2:MM3**, and **ON12:MM5** with **ON2:MM5**). In contrast, less efficient discrimination is observed when the mismatched base-pair is located close to the bulge (e.g., compare  $\Delta T_{ms}$  of **ON12:MM1** with **ON2:MM1**), presumably since the bulge region introduces greater flexibility in the duplex, which affects base-pairing fidelity. Along these lines, increases in bulge size – at least up to a saturation point – result in progressive deterioration of binding specificity (e.g., compare  $\Delta T_{ms}$  of **ON4/ON6/ON8/ON10/ON12/ON14:MM1**). Similar trends have been observed with ONs modified with other non-nucleotidic bulges.<sup>S2</sup>



**Table S2.** Thermal denaturation temperatures of single mismatched DNA duplexes.<sup>a</sup>

ON	Sequence	$T_m$ [ $\Delta T_m$ ] (°C)		
		5'-ON:MM2 3'-ON:MM1	5'-ON:MM4 3'-ON:MM3	5'-ON:MM6 3'-ON:MM5
1:2	5'-GG <b>I</b> GGTCAACTATCT <b>I</b> GGA 3'-CCACCAGTTGATAGAC <b>C</b> CT	56.5 [-9.0] 63.0 [-5.5]	62.5 [-3.0] 61.5 [-7.0]	59.5 [-6.0] 59.5 [-9.0]
3:4	5'-GG <b>I</b> GGTCAA <b>2</b> CTATCT <b>I</b> GGA 3'-CCACCAGTT <b>2</b> GATAGAC <b>C</b> CT	53.5 [-9.5] 59.5 [-5.0]	58.0 [-5.0] 56.5 [-8.0]	56.5 [-6.5] 54.5 [-10.0]
5:6	5'-GG <b>I</b> GGTCAA <b>222</b> CTATCT <b>I</b> GGA 3'-CCACCAGTT <b>222</b> GATAGAC <b>C</b> CT	51.5 [-6.5] 55.0 [-4.0]	55.5 [-2.5] 51.0 [-8.0]	53.0 [-5.0] 48.0 [-11.0]
7:8	5'-GG <b>I</b> GGTCAA <b>4</b> CTATCT <b>I</b> GGA 3'-CCACCAGTT <b>4</b> GATAGAC <b>C</b> CT	51.0 [-7.0] 55.0 [-5.0]	54.5 [-3.5] 51.0 [-9.0]	52.0 [-6.0] 49.0 [-11.0]
9:10	5'-GG <b>I</b> GGTCAA <b>444</b> CTATCT <b>I</b> GGA 3'-CCACCAGTT <b>444</b> GATAGAC <b>C</b> CT	46.5 [-6.5] 51.0 [-2.5]	50.0 [-3.0] 44.0 [-9.5]	46.5 [-6.5] 43.0 [-10.5]
11:12	5'-GG <b>I</b> GGTCAA <b>9</b> CTATCT <b>I</b> GGA 3'-CCACCAGTT <b>9</b> GATAGAC <b>C</b> CT	48.5 [-5.0] 53.5 [-2.0]	49.5 [-4.0] 47.0 [-8.5]	46.5 [-7.0] 44.0 [-11.5]
13:14	5'-GG <b>I</b> GGTCAA <b>999</b> CTATCT <b>I</b> GGA 3'-CCACCAGTT <b>999</b> GATAGAC <b>C</b> CT	45.0 [-8.0] 54.0 [-2.0]	47.0 [-6.0] 47.5 [-8.5]	48.5 [-4.5] 47.0 [-9.0]
DNA1	5'-GGTGGTCAACTATCTGGA	50.5 [-8.5]	55.5 [-3.5]	50.5 [-8.5]
DNA2	3'-CCACCAGTTGATAGACCT	53.5 [-5.5]	51.5 [-7.5]	47.5 [-11.5]

<sup>a</sup> For experimental conditions, see Table 1.  $\Delta T_m$  is calculated relative to the  $T_m$  of the duplexes between the specific probe strand (ON) and its complementary DNA (values listed in Table 1). **MM1**: 5'-GGTGGTCAGCTATCTGGA; **MM2**: 3'-CCACCAGTCGATAGACCT; **MM3**: 5'-GGTGGCCA ACTATCTGGA; **MM4**: 3'-CCACCGGTTGATAGACCT; **MM5**: 5'-GGTAGTCAACTATCTGGA; **MM6**: 3'-CCATCAGTTGATAGACCT. Mismatched nucleotides relative to the probe sequence are italicized.

**Table S3.** Thermal advantage ( $TA$ ) values for different Invader probes.<sup>a</sup>

Sequence			
5'-GG <u>U</u> GGTCAA X <sub>1</sub> CTATC <u>U</u> GGA			
3'-CCAC <u>C</u> CAGTT X <sub>2</sub> GATAGAC <u>C</u> CT			
Probe	X <sub>1</sub>	X <sub>2</sub>	$TA$ (°C)
1:2	-	-	+14.5
3:2	2	-	+14.5
1:4	-	2	+17.0
3:4	2	2	+16.5
5:2	222	-	+15.5
1:6	-	222	+18.0
5:6	222	222	+17.0
7:2	4	-	+16.0
1:8	-	4	+19.0
7:8	4	4	+18.5
9:2	444	-	+18.5
1:10	-	444	+20.5
9:10	444	444	+10.5
11:2	9	-	+19.0
1:12	-	9	+20.0
11:12	9	9	+6.0
13:2	999	-	+17.5
1:14	-	999	+18.0
13:14	999	999	+5.0

<sup>a</sup>  $TA = T_m (5\text{'-ON:cDNA}) + T_m (3\text{'-ON:cDNA}) - T_m (\text{Invader probe}) - T_m (\text{dsDNA})$ .  $T_m$  values underlying the calculations are shown in Table 1.

*Additional discussion of the  $\Delta G_{rec}$  term.* The term  $\Delta G_{rec}$ , which we define as  $\Delta G_{rec} = \Delta G_{5'-ON:cDNA} + \Delta G_{3'-ON:cDNA} - \Delta G_{Invader\ probe} - \Delta G_{dsDNA}$ , describes the thermodynamic driving force at a given temperature for Invader-mediated recognition of an isosequential dsDNA target (i.e., 5'-GGTGGTCAACTATCTGGA : 3'-CCACCAGTTGATAGACCT for the Invader probes listed in Table 1). While  $\Delta G_{rec}$  does not provide the thermodynamic driving force for recognition of complex targets such as **DH1** or chromosomal DNA (*vide supra*), it is a useful term to assess different Invader designs relative to each other. The  $\Delta G_{dsDNA}$  component – i.e., the change in free energy associated with formation of the dsDNA target region – is likely more negative (i.e., more stable) for complex targets such as **DH1** or chromosomal DNA regions since several of the base pairs surrounding the targeted region likely must be broken to accommodate the double-duplex invasion complex that is presumed formed as part of the recognition process. The  $\Delta G_{5'-ON:cDNA}$  and  $\Delta G_{3'-ON:cDNA}$  terms are not expected to deviate substantially in the formed recognition complexes except for potential stacking/capping interactions with neighboring nucleotides. Consequently, we expect  $\Delta G_{rec}$  values for Invader-mediated recognition of complex targets to be considerably larger (less favorable) than for recognition of isosequential dsDNA targets. This is supported by our experimental observations (e.g., **ON13:ON14** only results in trace recognition of **DH1**, Figure 2, despite a  $\Delta G_{rec}^{298}$  value of -15 kJ/mol, Table 1). Equivalent considerations apply to the  $T_m$ -based TA term.

*Additional discussion of thermodynamic parameters.* The trends in Gibbs free energy associated with duplex formation at 298 K (Table S4) generally reinforce the observed  $T_m$  trends that are discussed at length in the main manuscript: 1) Formation of the conventional Invader probe **ON1:ON2** is thermodynamically less favorable than formation of the corresponding unmodified

DNA duplex (i.e.,  $\Delta\Delta G^{298} = 14$  kJ/mol). 2) Insertion of bulges invariably further decreases the stability of the double-stranded probes (i.e.,  $\Delta\Delta G^{298}$  between 26-45 kJ/mol). 3) The stability of Invader probes with a single bulge decreases as the size of the bulge increases (e.g.,  $\Delta G^{298} = -71$  kJ/mol,  $-69$  kJ/mol, and  $-58$  kJ/mol, for **ON3:ON2**, **ON7:ON2**, and **ON11:ON2**, respectively). 4) Expansion of the non-nucleotidic region by incorporation of three consecutive bulge monomers results in further destabilization of single-bulge Invader probes when small, but not large, monomers are used (e.g., compare  $\Delta G^{298}$  values for **ON3:ON2** and **ON5:ON2**, **ON7:ON2** and **ON9:ON2**, and **ON11:ON2** and **ON13:ON2**). 5) Invader probes with two opposing bulges comprised of **2**, **4**, or **222** segments are additionally, but not additively, destabilized relative to the corresponding single-bulge probes, whereas probes with two opposing bulges comprised of **9**, **444**, or **999** segments are not additionally destabilized (e.g., compare  $\Delta G^{298}$  values for **ON7:ON2**, **ON1:ON8**, and **ON7:ON8** with **ON11:ON2**, **ON1:ON12**, and **ON11:ON12**). 6) Duplexes between individual conventional probe strands and cDNA are more stable than the corresponding unmodified DNA duplex (i.e.,  $\Delta\Delta G^{298} = -10$  kJ/mol and  $-15$  kJ/mol for **ON1:cDNA** and **ON2:cDNA**, respectively). 7) Incorporation of a bulge decreases the stability of duplexes between individual probe strands and cDNA (i.e.,  $\Delta\Delta G^{298} = 0-18$  kJ/mol). Incorporation of large bulges generally results in greater destabilization, although the trends are not fully consistent (e.g., compare  $\Delta\Delta G^{298}$  values for duplexes between **ON3**, **ON7**, or **ON11** and cDNA). 8) There are no clear monomer-specific trends as it pertains to the destabilizing effects of duplexes between probe strands and cDNA upon expansion of the bulges (e.g., compare  $\Delta G^{298}$  values for **ON3:cDNA** and **ON5:cDNA** with **ON4:cDNA** and **ON6:cDNA**).

As discussed in the main manuscript (Table 1), the combined impact of these trends on the driving force for Invader-mediated dsDNA-recognition at 298 K ( $\Delta G_{\text{rec}}^{298}$ ) are: 1) Probes featuring

a single **2** or **4** bulge display slightly lower thermodynamic driving forces for dsDNA-recognition than conventional probe **ON1:ON2**. 2) Probes featuring a single **222**, **444**, **9**, or **999** bulge display driving forces of similar or greater magnitude than the conventional Invader probe. 3) Probes with two opposing bulges generally display far less favorable thermodynamic driving forces.

The relative  $\Delta\Delta G$  trends for formation of probe duplexes or duplexes between individual probes strands and cDNA at 310 K are similar. Thus, probe duplexes are  $\sim 8$  kJ/mol more stable, whilst duplexes between individual probes strands and cDNA are  $\sim 2$  kJ/mol more stable relative to the unmodified reference duplex at 310 K than at 298 K (compare  $\Delta\Delta G$  values for duplexes at 298 K and 310 K, Table S4 and S5). The combined effect of these trends is that  $\Delta G_{\text{rec}}^{310}$  values, on average, are  $\sim 4$  kJ/mol less favorable (compare  $\Delta G_{\text{rec}}^{298}$  and  $\Delta G_{\text{rec}}^{310}$ , Table 1).

The van't Hoff method generally yields robust datasets for  $\Delta G$  values, whereas  $\Delta H$  and  $\Delta S$  values are highly sensitive to the choice and fitting of baselines.<sup>S3</sup> Accordingly, we acknowledge the following trends in  $\Delta H$  and  $-T\Delta S$  values but caution against overinterpretation of the results (Tables S6 and S7): 1) Formation of conventional Invader probe **ON1:ON2** is far less enthalpically favorable than formation of the corresponding unmodified DNA duplex (i.e.,  $\Delta\Delta H = +152$  kJ/mol). Thus, enthalpic factors largely account for the increased lability of **ON1:ON2**. This is in agreement with the expectation that base-pairing is perturbed near the +1 interstrand zipper arrangements of 2'-*O*-(pyren-1-yl)methyl RNA monomers as the neighbor exclusion principle is violated. 2) Incorporation of bulges into Invader probes invariably results in additional enthalpic penalties, presumably as nearby base-pairs are weakened due to the introduced flexibility (i.e.,  $\Delta\Delta H = 166$ - $276$  kJ/mol for bulge-containing Invader probes). 3) The  $\Delta H$  values for Invader probes with a single incorporation of monomers **2**, **4**, or **9** not vary in a size-dependent manner (e.g., compare  $\Delta H$  for **ON3:ON2**, **ON7:ON2**, and **ON11:ON2**) as the presence of monomer **4** results in the least

pronounced enthalpic penalties. 4) The impact on  $\Delta H$  values of single-bulge probes upon expansion of the bulge through incorporation of consecutive **2**, **4**, or **9** monomers is limited and does not appear to follow a clear monomer-specific trend (change in  $\Delta H$  ranging from additional destabilization by  $\sim 50$  kJ/mol to stabilization by  $\sim 40$  kJ/mol; e.g., compare  $\Delta H$  for **ON3:ON2** and **ON5:ON2**, and **ON1:ON4** and **ON1:ON6**). 5) Addition of a second bulge results in minor additional enthalpic destabilization of probe duplexes when small linker monomers are used (i.e., monomers **2** or **4**). No additional enthalpic destabilization of Invader probes is observed upon incorporation of a second larger bulge (i.e., **222**, **444**, **9** or **999** segments). 6) As discussed earlier, duplexes between conventional probe strands and cDNA are more thermodynamically stable than the corresponding unmodified DNA duplex. Duplex formation is accompanied by a less unfavorable change in entropy (e.g.,  $\Delta(T\Delta S^{298}) = -14$  kJ/mol for **ON1:cDNA**). The results suggest that the 2'-*O*-(pyren-1-yl)methyl RNA monomers may preorganize probe strands for duplex formation with cDNA (e.g., through stacking between the pyrene moiety and proximal nucleobases) but other factors, including hydration effects, likely also contribute. 7) With few exceptions, incorporation of bulges renders duplex formation between probe strands and cDNA enthalpically less favorable ( $\Delta\Delta H$  values between  $-4$  and  $+167$  kJ/mol). There are no clear bulge size-dependent trends on the  $\Delta H$  values of probe-cDNA duplexes. 9) Expansion of the bulges through incorporation of consecutive non-nucleotidic monomers appears to introduce significant entropic destabilization of the probe:cDNA duplexes, presumably reflecting increased flexibility (e.g., compare  $\Delta(T\Delta S^{298})$  values for **ON3:cDNA** and **ON5:cDNA**).

The overall impact of these enthalpy and entropy trends are: 1) The thermodynamically favorable driving force for Invader-mediated recognition of isosequential dsDNA is strongly enthalpy-driven (i.e.,  $\Delta H_{\text{rec}} \ll 0$  kJ/mol, Table S6). 2) The very favorable driving forces observed

for Invader probes with a single large bulge (i.e., featuring a **222**, **444**, **9**, or **999** segment) are due to particularly favorable overall changes in enthalpy ( $\Delta H_{\text{rec}}$  between -266 and -136 kJ/mol, Table S6), which are partially counterbalanced by overall entropic penalties (Table S7 and S8). 3) The less favorable driving forces for dsDNA-recognition seen for Invader probes with two adjacent bulges are due to less favorable changes in enthalpy ( $\Delta H_{\text{rec}} > -95$  kJ/mol, Table S6). Invader probe **ON5:ON6** with two adjacent **222** segments is an exception hereto as it displays favorable  $\Delta G_{\text{rec}}$  values and a very favorable  $\Delta H_{\text{rec}}$  value of -224 kJ/mol, which, in largest part, ensues because formation of the probe duplex is associated with a very prominent enthalpic penalty (compare  $\Delta H_{\text{rec}}$  for **ON5:ON6** and **ON1:ON2**, Table S6). The underlying reasons for this result are not fully understood.

**Table S4.** Change in Gibbs free energy at 298 K ( $\Delta G^{298}$ ) upon formation of double-stranded probes or duplexes between individual probe strands and cDNA. Also shown is the calculated change in free energy upon probe-mediated recognition of isosequential dsDNA targets ( $\Delta G_{\text{rec}}^{298}$ ).<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACCT			$\Delta G^{298}$ [ $\Delta\Delta G^{298}$ ] (kJ/mol)			
ON	X <sub>1</sub>	X <sub>2</sub>	Invader Probe	5'-ON:cDNA	3'-ON:cDNA	$\Delta G_{\text{rec}}^{298}$ (kJ/mol)
1:2	-	-	-83 [+14]	-107 [-10]	-112 [-15]	-39
3:2	2	-	-71 [+26]	-83 [+14]	-112 [-15]	-27
1:4	-	2	-70 [+27]	-107 [-10]	-95 [+2]	-35
3:4	2	2	-64 [+33]	-83 [+14]	-95 [+2]	-17
5:2	222	-	-69 [+28]	-97 [ $\pm$ 0]	-112 [-15]	-43
1:6	-	222	-59 [+38]	-107 [-10]	-93 [+4]	-44
5:6	222	222	-54 [+43]	-97 [ $\pm$ 0]	-93 [+4]	-39
7:2	4	-	-69 [+28]	-90 [+7]	-112 [-15]	-36
1:8	-	4	-64 [+33]	-107 [-10]	-82 [+15]	-28
7:8	4	4	-52 [+45]	-90 [+7]	-82 [+15]	-23
9:2	444	-	-61 [+36]	-84 [+13]	-112 [-15]	-38
1:10	-	444	-52 [+45]	-107 [-10]	-83 [+14]	-41
9:10	444	444	-53 [+44]	-84 [+13]	-83 [+14]	-17
11:2	9	-	-58 [+39]	-79 [+18]	-112 [-15]	-36
1:12	-	9	-54 [+43]	-107 [-10]	-84 [+13]	-40
11:12	9	9	-56 [+41]	-79 [+18]	-84 [+13]	-10
13:2	999	-	-59 [+38]	-80 [+17]	-112 [-15]	-36
1:14	-	999	-58 [+39]	-107 [-10]	-95 [+2]	-47
13:14	999	999	-63 [+34]	-80 [+17]	-95 [+2]	-15

<sup>a</sup>  $\Delta\Delta G^{298}$  is calculated relative to the corresponding unmodified DNA duplex ( $\Delta G^{298} = -97$  kJ/mol).  
 $\Delta G_{\text{rec}}^{298} = \Delta G^{298} (5\text{'-ON:cDNA}) + \Delta G^{298} (3\text{'-ON:cDNA}) - \Delta G^{298} (\text{Invader probe}) - \Delta G^{298} (\text{dsDNA})$ .



**Table S5.** Change in Gibbs free energy at 310 K ( $\Delta G^{310}$ ) upon formation of double-stranded probes or duplexes between individual probe strands and cDNA. Also shown is the calculated change in free energy upon probe-mediated recognition of isosequential dsDNA targets ( $\Delta G_{\text{rec}}^{310}$ ).<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACT			$\Delta G^{310}$ [ $\Delta\Delta G^{310}$ ] (kJ/mol)			$\Delta G_{\text{rec}}^{310}$ (kJ/mol)
ON	X <sub>1</sub>	X <sub>2</sub>	Invader probe	5'-Inv: cDNA	3'-Inv: cDNA	
1:2	-	-	-69 [+8]	-88 [-11]	-93 [-16]	-35
3:2	2	-	-60 [+17]	-70 [+7]	-93 [-16]	-26
1:4	-	2	-58 [+19]	-88 [-11]	-78 [-1]	-31
3:4	2	2	-53 [+24]	-70 [+7]	-78 [-1]	-18
5:2	222	-	-56 [+21]	-74 [+3]	-93 [-16]	-34
1:6	-	222	-49 [+28]	-88 [-11]	-77 [ $\pm$ 0]	-39
5:6	222	222	-42 [+35]	-74 [+3]	-77 [ $\pm$ 0]	-32
7:2	4	-	-55 [+22]	-72 [+5]	-93 [-16]	-33
1:8	-	4	-51 [+26]	-88 [-11]	-68 [+9]	-28
7:8	4	4	-41 [+36]	-72 [+5]	-68 [+9]	-22
9:2	444	-	-47 [+30]	-65 [+12]	-93 [-16]	-34
1:10	-	444	-41 [+36]	-88 [-11]	-65 [+12]	-35
9:10	444	444	-38 [+39]	-65 [+12]	-65 [+12]	-15
11:2	9	-	-46 [+31]	-62 [+15]	-93 [-16]	-32
1:12	-	9	-43 [+34]	-88 [-11]	-66 [+11]	-34
11:12	9	9	-45 [+32]	-62 [+15]	-66 [+11]	-6
13:2	999	-	-47 [+30]	-63 [+14]	-93 [-16]	-32
1:14	-	999	-46 [+31]	-88 [-11]	-75 [+2]	-40
13:14	999	999	-49 [+28]	-63 [+14]	-75 [+2]	-12

<sup>a</sup>  $\Delta\Delta G^{310}$  is calculated relative to the corresponding unmodified DNA duplex ( $\Delta G^{310} = -77$  kJ/mol).  
 $\Delta G_{\text{rec}}^{310} = \Delta G^{310} (5\text{'-ON:cDNA}) + \Delta G^{310} (3\text{'-ON:cDNA}) - \Delta G^{310} (\text{Invader probe}) - \Delta G^{310} (\text{dsDNA})$ .

**Table S6.** Change in enthalpy ( $\Delta H$ ) upon formation of double-stranded probes (Invader probe) or duplexes between individual probe strands and cDNA. Also shown is the calculated change in enthalpy upon Invader-mediated recognition of isosequential dsDNA targets ( $\Delta H_{rec}$ ).<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACCT			$\Delta H$ [ $\Delta\Delta H$ ] (kJ/mol)			
ON	X <sub>1</sub>	X <sub>2</sub>	Invader probe	5'-Inv: cDNA	3'-Inv: cDNA	$\Delta H_{rec}$ (kJ/mol)
1:2	-	-	-436 [+152]	-584 [+4]	-587 [+1]	-147
3:2	2	-	-345 [+243]	-421 [+167]	-587 [+1]	-75
1:4	-	2	-361 [+227]	-584 [+4]	-511 [+77]	-146
3:4	2	2	-338 [+250]	-421 [+167]	-511 [+77]	-6
5:2	222	-	-372 [+216]	-561 [+27]	-587 [+1]	-188
1:6	-	222	-322 [+266]	-584 [+4]	-592 [-4]	-266
5:6	222	222	-341 [+247]	-561 [+27]	-592 [-4]	-224
7:2	4	-	-406 [+182]	-535 [+53]	-587 [+1]	-128
1:8	-	4	-381 [+207]	-584 [+4]	-431 [+157]	-46
7:8	4	4	-312 [+276]	-535 [+53]	-431 [+157]	-66
9:2	444	-	-410 [+178]	-556 [+32]	-587 [+1]	-145
1:10	-	444	-328 [+260]	-584 [+4]	-536 [+52]	-204
9:10	444	444	-422 [+166]	-556 [+32]	-536 [+52]	-82
11:2	9	-	-354 [+234]	-495 [+93]	-587 [+1]	-140
1:12	-	9	-317 [+271]	-584 [+4]	-513 [+75]	-192
11:12	9	9	-337 [+251]	-495 [+93]	-513 [+75]	-83
13:2	999	-	-361 [+227]	-498 [+90]	-587 [+1]	-136
1:14	-	999	-356 [+232]	-584 [+4]	-600 [-12]	-240
13:14	999	999	-415 [+173]	-498 [+90]	-600 [-12]	-95

<sup>a</sup>  $\Delta\Delta H$  is calculated relative to the corresponding unmodified DNA duplex ( $\Delta H = -588$  kJ/mol).  
 $\Delta H_{rec} = \Delta H(5'-ON:cDNA) + \Delta H(3'-ON:cDNA) - \Delta H(\text{Invader probe}) - \Delta H(\text{dsDNA})$ .

**Table S7.** Change in entropy at 298 K ( $-T^{298}\Delta S$ ) upon formation of double-stranded probes or duplexes between individual probe strands and cDNA. Also shown is the calculated change in entropy upon Invader-mediated recognition of isosequential dsDNA targets ( $-T\Delta S_{\text{rec}}^{298}$ ).<sup>a</sup>

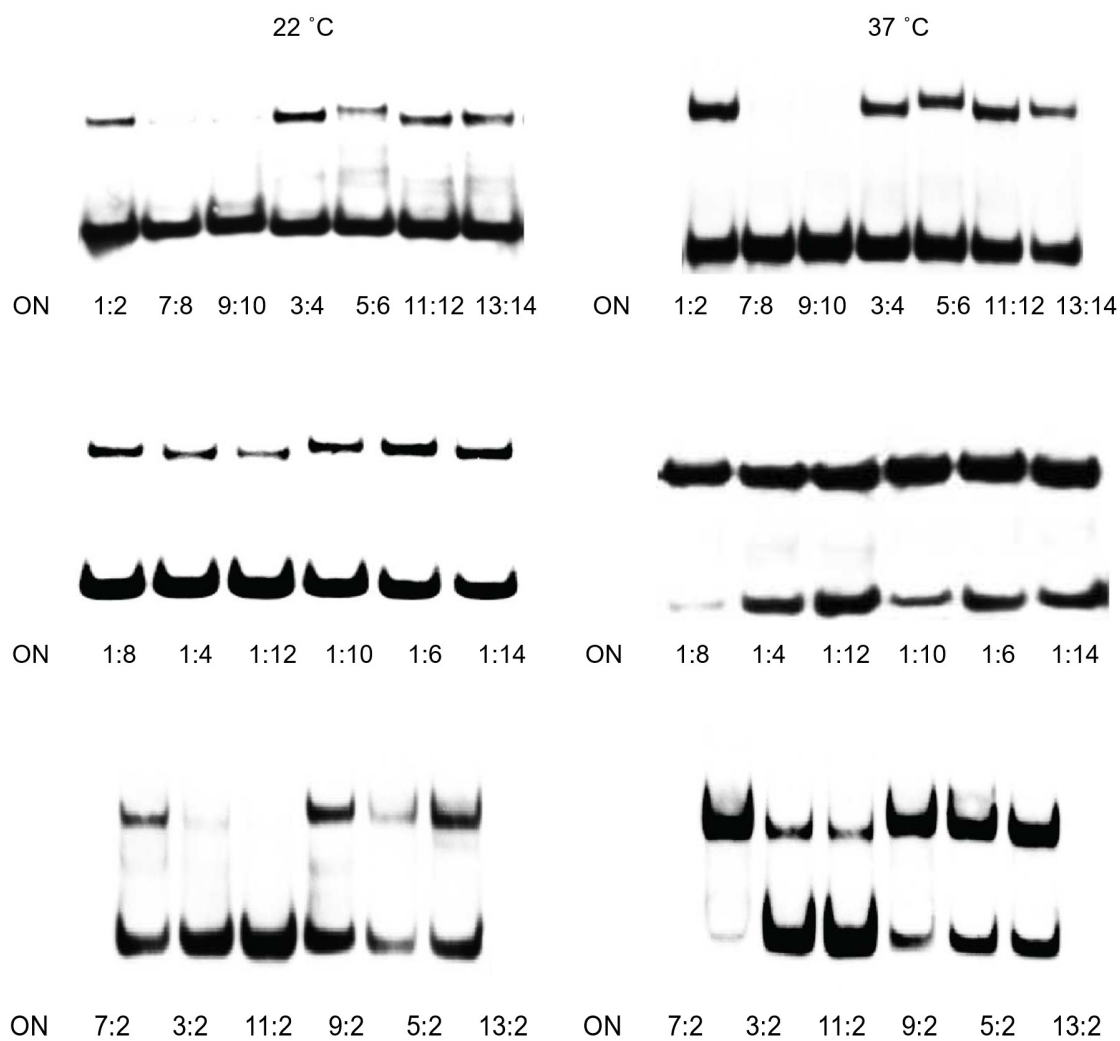
Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGGA 3'-CCACCAGTT X <sub>2</sub> GATAGACCT			$-T\Delta S^{298}$ [ $\Delta(T\Delta S^{298})$ ] (kJ/mol)			
ON	X <sub>1</sub>	X <sub>2</sub>	Invader probe	5'-Inv: cDNA	3'-Inv: cDNA	$-T\Delta S_{\text{rec}}^{298}$ (kJ/mol)
1:2	-	-	353 [-138]	477 [-14]	475 [-16]	108
3:2	2	-	274 [-217]	337 [-154]	475 [-16]	47
1:4	-	2	291 [-200]	477 [-14]	415 [-76]	110
3:4	2	2	274 [-217]	337 [-154]	415 [-76]	-13
5:2	222	-	304 [-187]	468 [-23]	475 [-16]	148
1:6	-	222	263 [-228]	477 [-14]	495 [+4]	218
5:6	222	222	287 [-204]	495 [+4]	468 [-23]	185
7:2	4	-	337 [-154]	445 [-46]	475 [-16]	92
1:8	-	4	317 [-174]	477 [-14]	349 [-142]	18
7:8	4	4	260 [-231]	445 [-46]	349 [-142]	43
9:2	444	-	349 [-142]	472 [-19]	475 [-16]	107
1:10	-	444	275 [-216]	477 [-14]	453 [-38]	164
9:10	444	444	369 [-122]	472 [-19]	453 [-38]	65
11:2	9	-	297 [-194]	416 [-75]	475 [-16]	103
1:12	-	9	263 [-228]	477 [-14]	429 [-62]	152
11:12	9	9	280 [-211]	416 [-75]	429 [-62]	74
13:2	999	-	302 [-189]	418 [-73]	475 [-16]	100
1:14	-	999	298 [-193]	477 [-14]	505 [+14]	193
13:14	999	999	352 [-139]	418 [-73]	505 [+14]	80

<sup>a</sup>  $\Delta(T\Delta S^{298})$  is calculated relative to the corresponding unmodified DNA duplex ( $-T\Delta S^{298} = 491$  kJ/mol).  $-T\Delta S_{\text{rec}}^{298} = \Delta(T\Delta S^{298})$  (5'-ON:cDNA) +  $\Delta(T\Delta S^{298})$  (3'-ON:cDNA) -  $\Delta(T\Delta S^{298})$  (Invader probe).

**Table S8.** Change in entropy at 310 K ( $-T\Delta S^{310}$ ) upon formation of double-stranded probes and duplexes between individual probe strands and cDNA. Also shown is the calculated change in entropy upon Invader-mediated recognition of isosequential dsDNA targets ( $-T\Delta S_{\text{rec}}^{310}$ ).<sup>a</sup>

Sequence						
5'-GGTGGTCAA X <sub>1</sub> CTATCTGGA						
3'-CCACCAGTT X <sub>2</sub> GATAGACT						
$-T\Delta S^{310}$ [ $\Delta(T\Delta S^{310})$ ] (kJ/mol)						
ON	X <sub>1</sub>	X <sub>2</sub>	Invader probe	5'-Inv: cDNA	3'-Inv: cDNA	$-T\Delta S_{\text{rec}}^{310}$ (kJ/mol)
1:2	-	-	367 [-144]	496 [-15]	494 [-17]	112
3:2	2	-	285 [-226]	351 [-160]	494 [-17]	49
1:4	-	2	303 [-208]	496 [-15]	432 [-79]	114
3:4	2	2	286 [-225]	351 [-160]	432 [-79]	-14
5:2	222	-	316 [-196]	487 [-24]	494 [-17]	154
1:6	-	222	273 [-238]	496 [15]	515 [+4]	227
5:6	222	222	298 [-213]	487 [-24]	515 [+4]	193
7:2	4	-	350 [-161]	463 [-48]	494 [-17]	96
1:8	-	4	330 [-181]	496 [-15]	363 [-148]	18
7:8	4	4	270 [-241]	463 [-48]	363 [-148]	45
9:2	444	-	363 [-148]	491 [-20]	494 [-17]	111
1:10	-	444	286 [-225]	496 [-15]	471 [-40]	170
9:10	444	444	383 [-128]	491 [-20]	471 [-40]	68
11:2	9	-	308 [-203]	432 [-79]	494 [-17]	107
1:12	-	9	274 [-237]	496 [-15]	447 [-64]	158
11:12	9	9	291 [-220]	432 [-79]	447 [-64]	77
13:2	999	-	314 [-197]	435 [-76]	494 [-17]	104
1:14	-	999	310 [-201]	496 [-15]	525 [+14]	200
13:14	999	999	366 [-145]	435 [-76]	525 [+14]	83

<sup>a</sup>  $\Delta(T\Delta S^{310})$  is calculated relative to the corresponding unmodified DNA duplex ( $-T\Delta S^{310} = 511$  kJ/mol).  $-T\Delta S_{\text{rec}}^{310} = \Delta(T\Delta S^{310})$  (5'-ON:cDNA) +  $\Delta(T\Delta S^{310})$  (3'-ON:cDNA) -  $\Delta(T\Delta S^{310})$  (Invader probe).

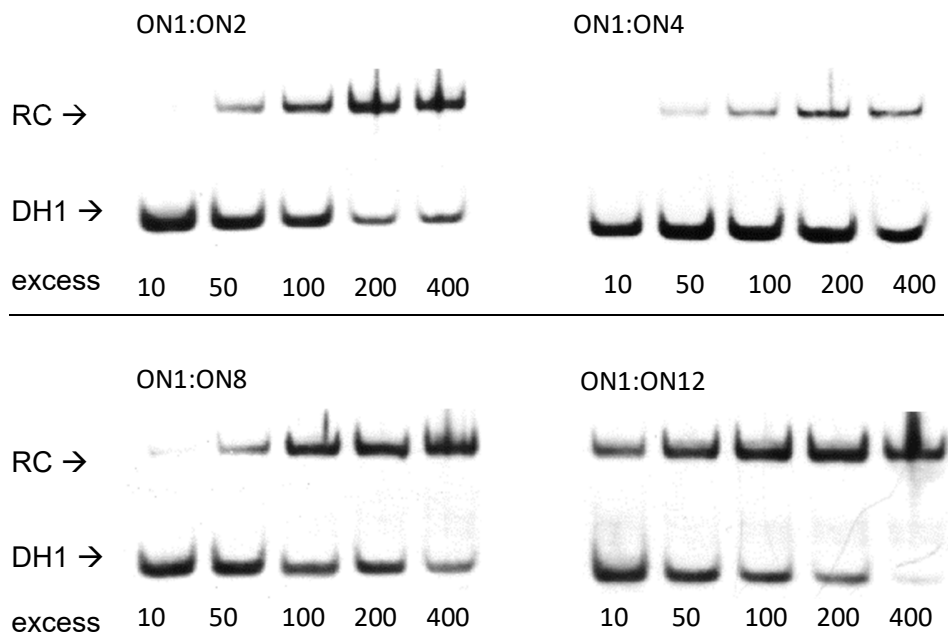


**Figure S6.** Representative electrophoretograms for recognition of DNA hairpin **DH1** using different Invader probes at room temperature (left) or 37 °C (right). For experimental conditions see Figure 2. For quantification of electrophoretograms, see Figure 2 and Table S9.

**Table S9.** Degree of **DH1**-recognition using 200-fold molar excess of different Invader probes.<sup>a</sup>





ON	Sequence	DH1-recognition (%)	
		22 °C	37 °C
1:2	5'-GG <b>I</b> GGTCAACTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTTGATAG <b>A</b> CCT	21 ± 6	60 ± 9
3:2	5'-GG <b>I</b> GGTCAA <b>2</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT GATAG <b>A</b> CCT	14 ± 10	16 ± 6
1:4	5'-GG <b>I</b> GGTCAA CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>2</b> GATAG <b>A</b> CCT	24 ± 5	51 ± 14
3:4	5'-GG <b>I</b> GGTCAA <b>2</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>2</b> GATAG <b>A</b> CCT	22 ± 6	51 ± 14
5:2	5'-GG <b>I</b> GGTCAA <b>222</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT GATAG <b>A</b> CCT	57 ± 8	65 ± 4
1:6	5'-GG <b>I</b> GGTCAA CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>222</b> GATAG <b>A</b> CCT	42 ± 6	64 ± 4
5:6	5'-GG <b>I</b> GGTCAA <b>222</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>222</b> GATAG <b>A</b> CCT	25 ± 7	40 ± 8
7:2	5'-GG <b>I</b> GGTCAA <b>4</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT GATAG <b>A</b> CCT	16 ± 7	23 ± 9
1:8	5'-GG <b>I</b> GGTCAA CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>4</b> GATAG <b>A</b> CCT	30 ± 9	61 ± 6
7:8	5'-GG <b>I</b> GGTCAA <b>4</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>4</b> GATAG <b>A</b> CCT	33 ± 7	48 ± 12
9:2	5'-GG <b>I</b> GGTCAA <b>444</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT GATAG <b>A</b> CCT	48 ± 6	74 ± 9
1:10	5'-GG <b>I</b> GGTCAA CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>444</b> GATAG <b>A</b> CCT	35 ± 5	70 ± 9
9:10	5'-GG <b>I</b> GGTCAA <b>444</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>444</b> GATAG <b>A</b> CCT	28 ± 1	40 ± 4
11:2	5'-GG <b>I</b> GGTCAA <b>9</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT GATAG <b>A</b> CCT	44 ± 6	89 ± 4
1:12	5'-GG <b>I</b> GGTCAA CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>9</b> GATAG <b>A</b> CCT	38 ± 9	82 ± 11
11:12	5'-GG <b>I</b> GGTCAA <b>9</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>9</b> GATAG <b>A</b> CCT	12 ± 4	10 ± 9
13:2	5'-GG <b>I</b> GGTCAA <b>999</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT GATAG <b>A</b> CCT	64 ± 24	76 ± 2
1:14	5'-GG <b>I</b> GGTCAA CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>999</b> GATAG <b>A</b> CCT	40 ± 11	77 ± 4
13:14	5'-GG <b>I</b> GGTCAA <b>999</b> CTATC <b>I</b> GGA 3'-CCAC <b>C</b> CAGTT <b>999</b> GATAG <b>A</b> CCT	8 ± 8	3 ± 7

<sup>a</sup> Quantification is based on the data shown in Figure 2.



**Figure S7.** Dose-response experiments. Representative electrophoretograms for recognition of model dsDNA target **DH1** (34.4 mM) using different concentrations of conventional and select bulge-containing Invader probes at 37 °C for 17 h. Experimental conditions are as described in Figure 2. For the corresponding dose-response curves, see Figure 3.

**Table S10.** Sequences and thermal denaturation temperatures of DNA hairpins used in the present study.<sup>a</sup>

DH	Sequence	$T_m$ (°C)
1	 5'-GGTGGTCAACTATCTGGA 3'-CCACCAGTTGATAGACCT	73.0
2	 5'-GGTGGTCAGCTATCTGGA 3'-CCACCAGTCGATAGACCT	74.5
3	 5'-GGTGGCCAACTATCTGGA 3'-CCACCGGTTGATAGACCT	74.5
4	 5'-GGTAGTCAACTATCTGGA 3'-CCATCAGTTGATAGACCT	68.5

<sup>a</sup> For experimental conditions, see Table 1. Base-pairs shown in red font denote the difference in sequence vis-à-vis the Invader probes.



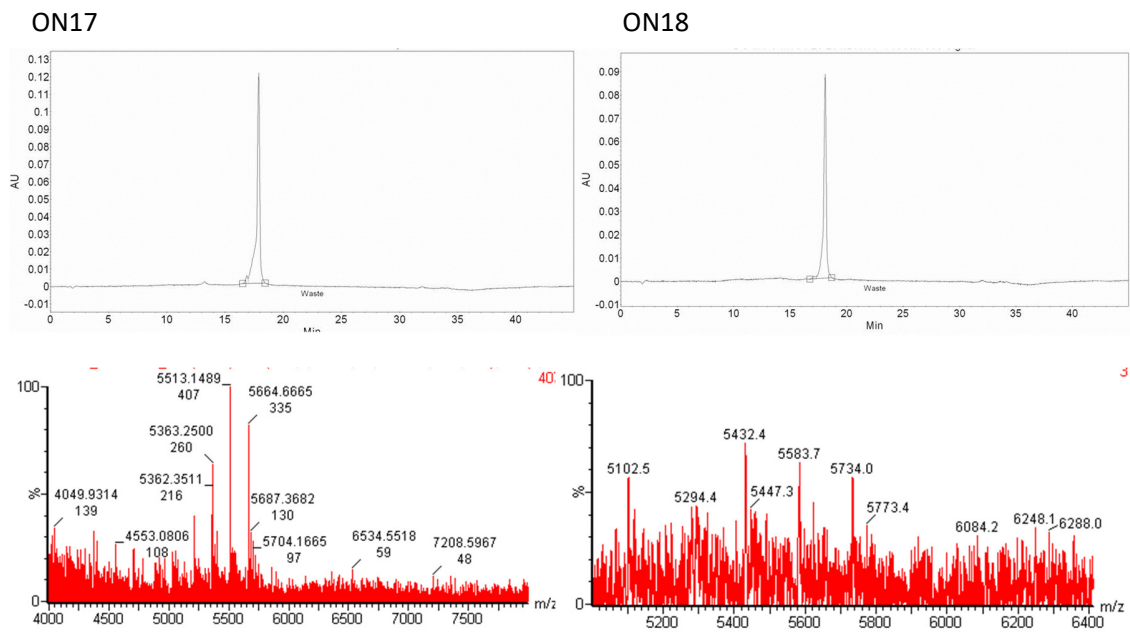
**Table S11.** MALDI-MS of *DYZ-1* targeting ONs.<sup>a</sup>

ON	Sequence	Observed <i>m/z</i> [M+H] <sup>+</sup>	Calculated <i>m/z</i> [M+H] <sup>+</sup>
15 <sup>b</sup>	5'-Cy3- <u>A</u> GCCC <u>U</u> GTGCC <u>C</u> TG	5397	5398
16 <sup>b</sup>	3'-TCGGGACACGGG <u>A</u> C-Cy3	5508	5510
17	3'-TCG GGACAC <u>4</u> GGG <u>A</u> C-Cy3	5665	5662
18	3'-TCGGGACAC <u>9</u> GGG <u>A</u> C-Cy3	5734	5732
19 <sup>c</sup>	5'-Cy3- <u>A</u> GCGC <u>U</u> GAGGC <u>C</u> TG	5486	5487
20	3'-TCG CGA <u>C</u> TC <u>4</u> CGG <u>A</u> C-Cy3	5576	5573
21	3'-TCG CGA <u>C</u> TC <u>9</u> CGG <u>A</u> C-Cy3	5646	5643
22	5'-Cy3- <u>A</u> GCCC <u>U</u> GTG <u>4</u> CC <u>C</u> TG	5553	5550
23	5'-Cy3- <u>A</u> GCCC <u>U</u> GTGCC <u>4</u> <u>C</u> TG	5552	5550
24	3'-TCGGGACACGG <u>4</u> G <u>A</u> C-Cy3	5663	5662
25	3'-TCGGGACACGG <u>9</u> G <u>A</u> C-Cy3	5734	5732
26	3'-TCG <u>9</u> GGACACGGG <u>A</u> C-Cy3	5735	5732

<sup>a</sup> Cy3 = Cy3 monomer, **4** = 4-hydroxybutyl monomer, **9** = 9-hydroxynonyl monomer, **A/U/C** = 2'-*O*-(pyren-1-yl)methyl RNA monomers. For structures, see Figure 1 in main manuscript. 2,4,6-trihydroxyacetophenone was used as the matrix.

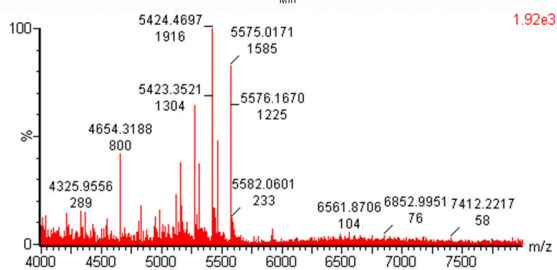
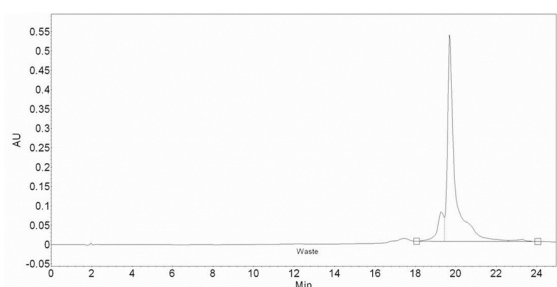
<sup>b</sup> Data previously published in Reference S4.

<sup>c</sup> Data previously published in Reference S5.

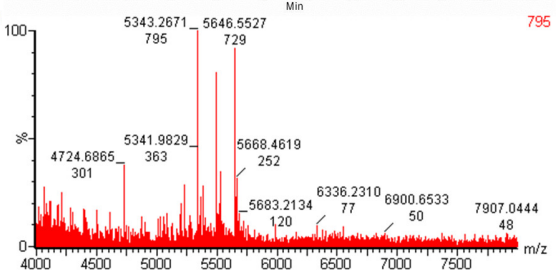
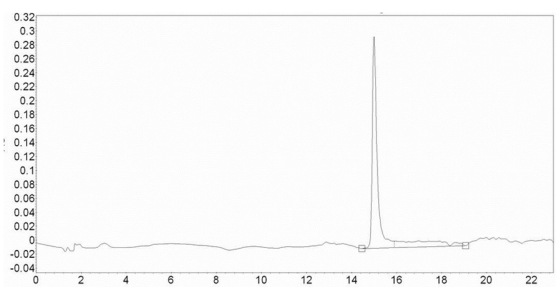


**Figure S8.** HPLC traces and MS spectra for **ON17** and **ON18**.

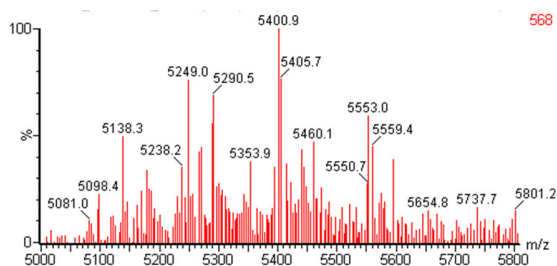
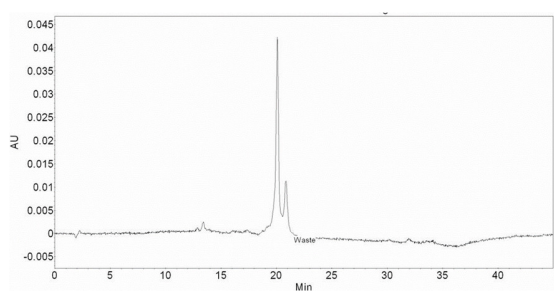
ON20



ON21



ON22



ON23

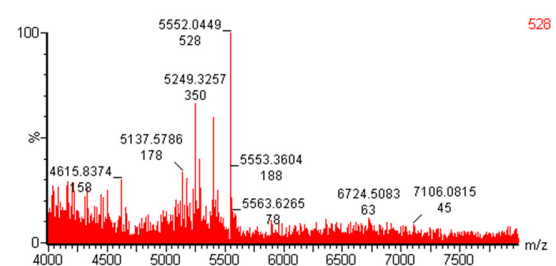
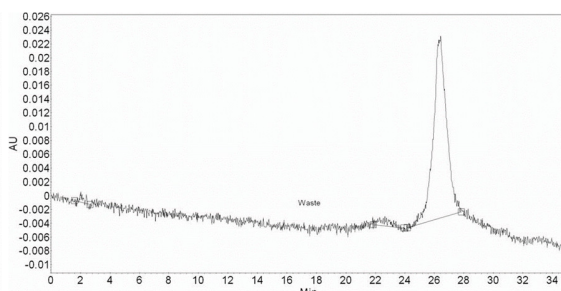
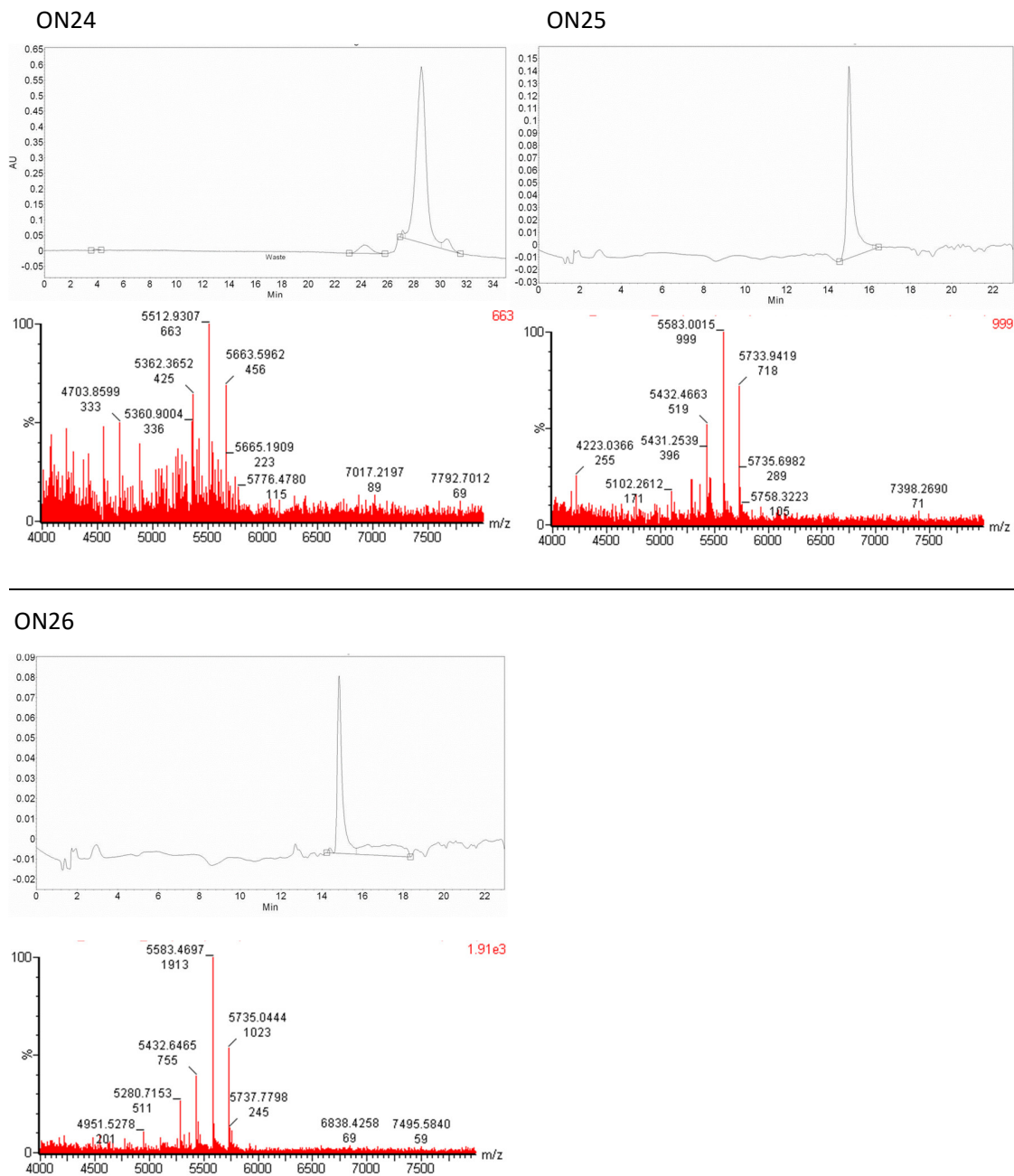
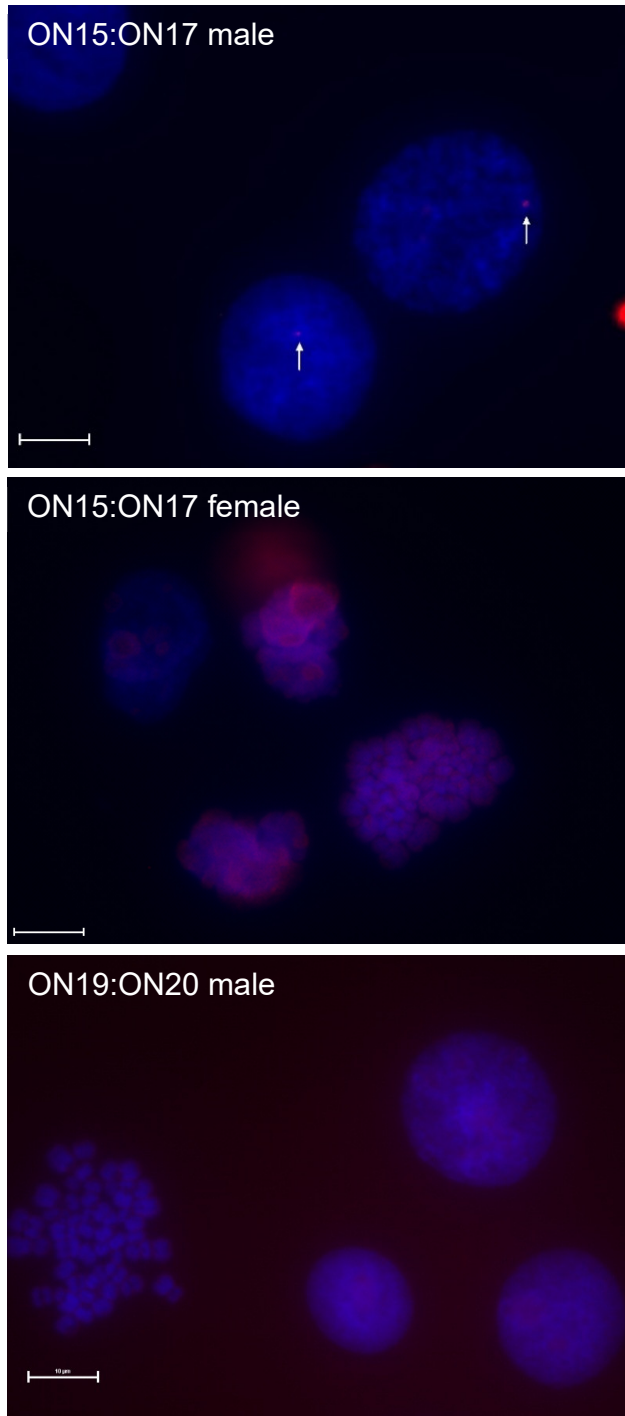


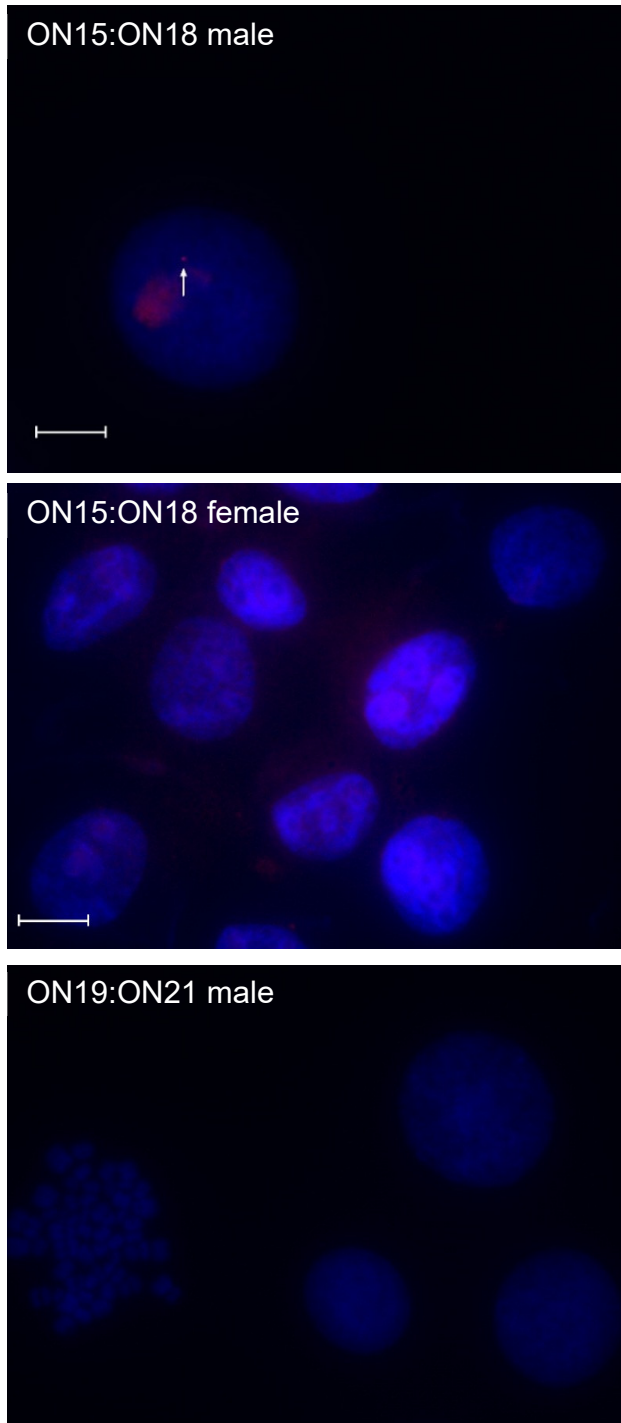
Figure S9. HPLC traces and MS spectra for ON20-ON23.



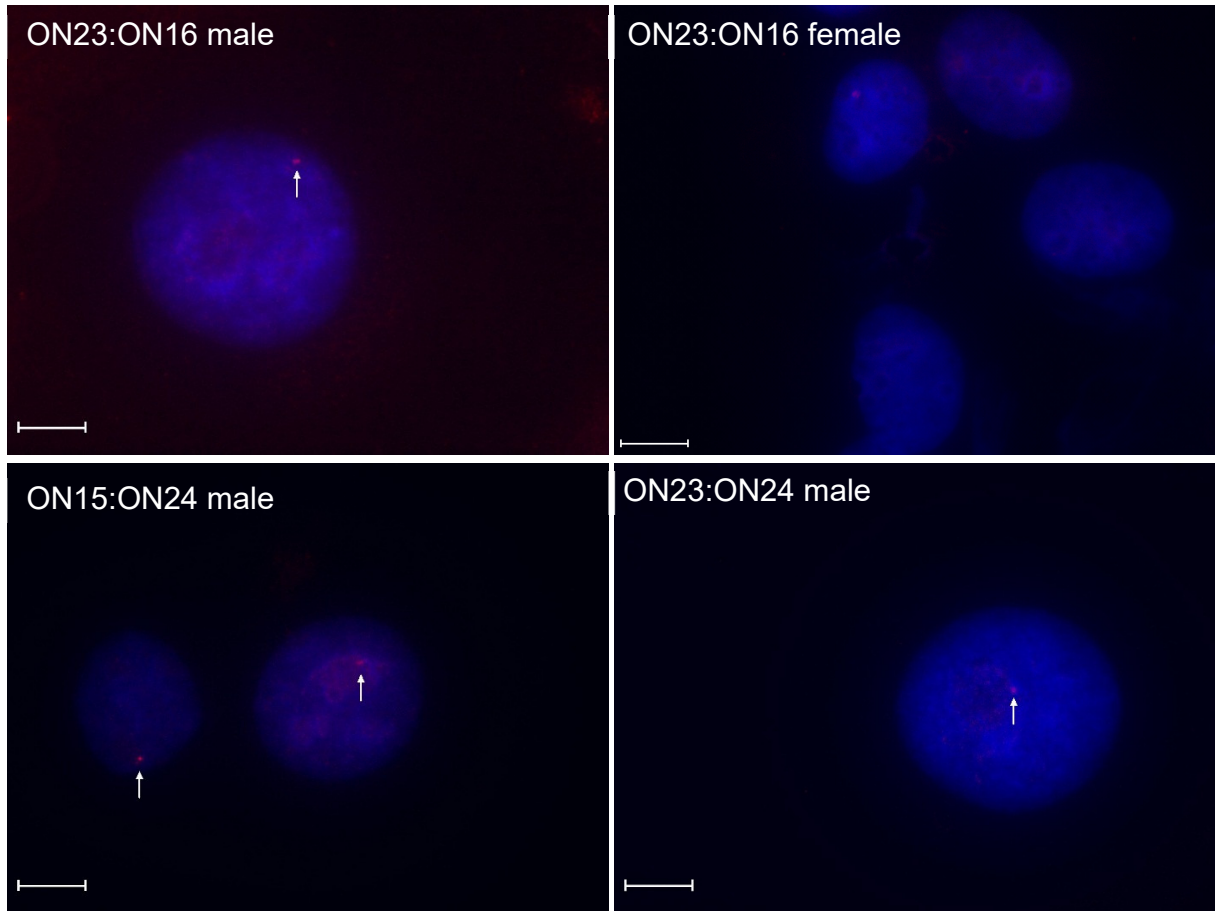
**Figure S10.** HPLC traces and MS spectra for ON24-ON26.



**Figure S11.** Representative images of Invader probe **ON15:ON17** or its triply mismatched control **ON19:ON20** incubated with nuclei from a male bovine kidney cell line (MDBK) or a female bovine endothelial cell line (CPAE) as indicated. For experimental details, see Figure 5.

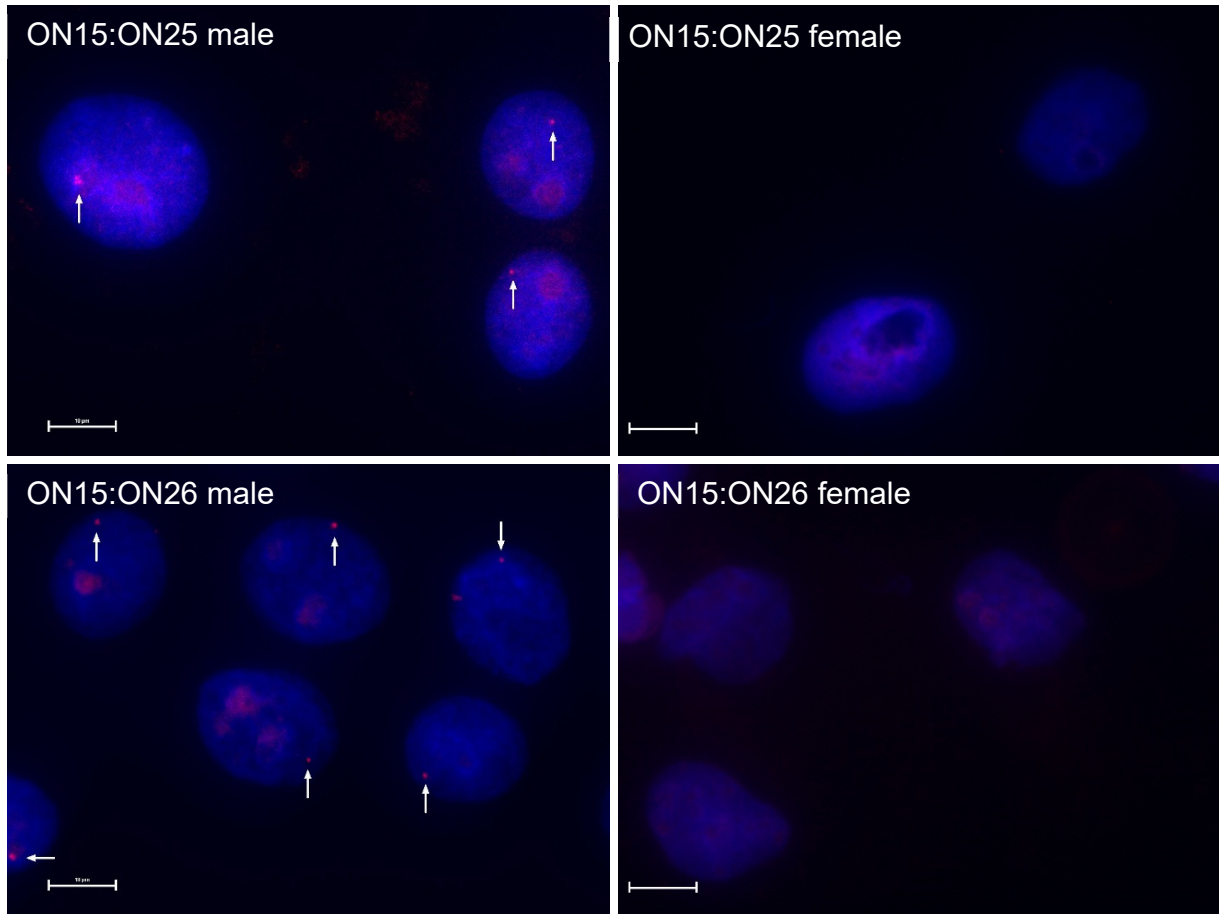


**Figure S12.** Representative images of Invader probe **ON15:ON18** or its triply mismatched control **ON19:ON21** incubated with nuclei from a male bovine kidney cell line (MDBK) or a female bovine endothelial cell line (CPAE) as indicated. For experimental details, see Figure 5.



**Figure S13.** Representative images of Invader probes **ON23:ON16**, **ON15:ON24**, and **ON23:ON24** incubated with nuclei from a male bovine kidney cell line (MDBK) or a female bovine endothelial cell line (CPAE) as indicated. For experimental details, see Figure 5.





**Figure S14.** Representative images of Invader probes **ON15:ON25** and **ON15:ON26** incubated with nuclei from a male bovine kidney cell line (MDBK) or a female bovine endothelial cell line (CPAE) as indicated. For experimental details, see Figure 5.



## SUPPLEMENTARY REFERENCES

S1) S. Karmakar, A. S. Madsen, D. C. Guenther, B. C. Gibbons and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2014, **12**, 7758-7773.

S2) N. Langkjær, A. Pasternak and J. Wengel, *Bioorg. Med. Chem. Lett.*, 2009, **17**, 5420-5425.

S3) J. L. Mergny and L. Lacroix, *Oligonucleotides*, 2003, **13**, 515-537.

S4) D. C. Guenther, G. H. Anderson, S. Karmakar, B. A. Anderson, B. A. Didion, W. Guo, J. P. Verstegen and P. J. Hrdlicka, *Chem. Sci.*, 2015, **6**, 5006-5015.

S5) B. A. Didion, S. Karmakar, D. C. Guenther, S. P. Sau, J. P. Verstegen and P. J. Hrdlicka, *ChemBioChem*, 2013, **14**, 1534-1538.